Research Paper Informatization Policy Vol. 33, No.1 (2026), pp.052-069
https://doi.org/10.22693/NIAIP.2026.33.1.052

프라이버시 규제와 디지털 담론의 구조적 전환: GDPR 전후 이용자 집단의 참여 구조 변화와 이슈-관심 주기 분석*

김주영**

요약 본 연구는 GDPR(General Data Protection Regulation) 시행이 온라인 프라이버시 담론의 참여 구조에 미친 사회기술적 영향을 실증적으로 규명하였다. Reddit의 r/privacy 커뮤니티에서 2016년부터 2023년까지 수집된 게시글(2,673건)과 댓글(38,944건)을 분석하여, 참여 양적 규모와 질적 차원 간 관계가 GDPR 전후 어떻게 변화했는지를 살펴보았다. 분석 결과, GDPR 이후 댓글 수의 유의한 증가는 관찰되지 않았으나, 댓글 점수는 참여 규모와 긍정적 상호작용을 보이며 질적 집중성이 강화되었다. 반면 댓글 길이는 발화가 간결화되는 패턴이 나타났다. 이러한 결과는 GDPR이 단순한 법 제도가 아니라 제도적(규제 시행), 기술적(플랫폼 어포던스 조정), 인식적(정보 공개 관리) 수준에서 담론의 구조를 재편한 사회기술적 사건으로 작용했음을 시사한다. 이론적으로 본 연구는 Downs(1972)의 이슈-관심 주기 이론을 디지털 규제 맥락에 확장하여, 사회적 관심의 상승-비용 인식-관심 감소 단계를 참여의 형식적 재편(formal restructuring) 과정으로 해석하였다. 실무적으로는 개인정보 규제가 참여의 양적 활성화보다는 질적 전환을 촉발함을 보여주며, 향후 데이터 거버넌스 및 플랫폼 설계 정책의 방향성을 제시한다.

주제어 : 프라이버시, GDPR, 디지털 담론, 댓글 분석, 이슈-관심 주기 분석

Privacy Regulation and Digital Discourse: The Issue-Attention Cycle of GDPR*

Kim, Jooyoung**

Abstract This study empirically examines the sociotechnical impact of the General Data Protection Regulation(GDPR) on participation structures within online privacy discourse. Using 2,673 posts and 38,944 comments collected from Reddit's r/privacy community between 2016 and 2023, the analysis investigates how the relationship between participation volume and quality changed before and after the implementation of GDPR. The results show no significant increase in comment volume after GDPR; however, comment scores exhibited a stronger positive association with participation volume, indicating intensified qualitative concentration, while comment length decreased, suggesting a shift toward more concise expressions. These findings suggest that GDPR functioned not merely as a legal framework but as a sociotechnical event that restructured online discourse across institutional, technical, and perceptual dimensions. Theoretically, the study extends Downs'(1972) issue-attention cycle to the digital regulation context, interpreting the stages of rising attention, cost recognition, and declining interest as a process of formal restructuring in participation. Practically, the findings demonstrate that privacy regulation triggers qualitative rather than quantitative transformation in online engagement, providing policy implications for future data-governance and platform-design strategies.

Keywords : privacy, GDPR, digital engagement, comment analysis, issue-attention cycle

접수: 2025. 9. 29; 최종수정: 2025. 10. 31; 게재확정: 2025. 11. 3.
* 이 논문은 한국연구재단의 지원을 받아 수행된 연구임(NRF-2025S1A5B5A17013726)
** 연세대학교 글로벌창의융합대학 경영학부 조교수(https://orcid.org/0000-0002-2256-7927)

I. 서론

디지털 시대의 도래와 함께 개인정보 보호는 단순한 법적 의무를 넘어 사회 전반에서 핵심적인 윤리적·기술적·정책적 의제로 부상하였다. 데이터 기반 사회에서 정보 주체의 권리 보장과 기업의 책임 있는 데이터 활용에 대한 요구가 커지면서, 유럽연합(EU)의 개인정보보호 일반 규정(General Data Protection Regulation, GDPR)은 이러한 변화를 가속화한 대표적인 제도적 전환점으로 평가된다(Rughinis et al., 2021). GDPR은 2018년 5월 시행 이후 전 세계 기업과 플랫폼에 강력한 데이터 수집·처리 규제를 도입하며, 개발자와 일반 사용자 모두가 법률적·기술적 대응을 고민하도록 만들었다. 실제로 GDPR은 온라인 포럼에서 법률적 지침 탐색을 증가시켰으며(Kyi et al., 2025), 일반 사용자와 개발자들이 Reddit과 같은 플랫폼에서 규제 해석과 실무적 대응 방안을 공유하도록 촉발하였다(Parsons et al., 2023). 이처럼 규제는 단순히 제도적 구속력을 넘어서, 온라인 커뮤니티 담론 자체의 주제와 참여 양상을 변화시키는 요인으로 작용한다.

동시에 온라인 플랫폼에서의 알고리즘 추천 시스템은 디지털 소비자의 일상적 경험을 지배하는 핵심 기술로 자리 잡았다. YouTube, TikTok, Netflix와 같은 대규모 추천 시스템은 사용자 데이터를 기반으로 콘텐츠를 선별·제시하며, 개인화 경험을 제공하는 동시에 체류 시간을 극대화한다(Covington et al., 2016). 그러나 이러한 설계는 본질적으로 플랫폼의 수익 극대화 논리에 맞추어져 있어, 참여의 양적 규모(volume)는 확대했으나 참여의 질적 차원(quality)을 약화시키거나 왜곡할 수 있다는 비판을 받아왔다(Siersdorfer et al., 2010; Zhou et al., 2010; Bishop, 2019). 즉, 플랫폼 참여 연구에서 양적 지표와 질적 지표 간 긴장 관계는 꾸준히 제기되어 왔으며, 이는 특히 정보시스템 및 디지털 정책 측면 모두에서 핵심적 학문적 쟁점으로 자리하고 있다.

GDPR의 규제 대상은 명목상 기업과 데이터처리자이지만, 그 효과는 이용자의 발화·참여 행위까지 확장되는 사회적 파급력을 지닌다. GDPR의 핵심 조항은 처리의 합법성(제6조), 정보 주체의 권리(제12조 이상), 자동화된 의사결정에 대한 거부권(제22조)을 규정하며, 특히 삭제권(right to erasure), 데이터 최소화(data minimization), 투명성(transparency) 의무는 플랫폼의 정책 설계와 사용자 인식 모두를 변화시킨다. GDPR 제20조에서 규정한 데이터 이동권은 정보 주체의 통제권 강화를 상징하며, 이용자 중심의 데이터 생태계로의 전환을 촉진했다(이창범 2021). 이러한 이용자 주권 중심의 규제 패러다임은 Reddit과 같은 온라인 커뮤니티에서도 발화 통제 및 자기노출 행태에 영향을 미칠 수 있다. 예컨대 Reddit, YouTube, Twitter 등 주요 플랫폼은 GDPR 시행 이후 개인정보 처리방침을 개정하고, 데이터 보존·삭제·동의 절차를 재정비하는 등 정책적 조정을 단행하였다(Goldberg et al., 2020). 이러한 기술적·정책적 변화는 이용자에게 자신의 발화가 더 오래 남거나 추적될 수 있다는 인식을 강화시켜, 발화의 길이·노출 범위·내용 구성 방식 등 미시적 행위에까지 영향을 미칠 수 있다(Petronio, 2002). Prasad & Pérez, 2020에 따르면 GDPR은 데이터 경제 전반에 제도적 충격을 주었으며, 단기적 비용 증가는 존재하지만 장기적으로 신뢰 기반의 디지털 생태계 조성을 촉진했다는 평가가 있다. 이는 규제가 단순한 제약이 아니라, 온라인 담론의 신뢰성과 질적 구조를 재편하는 매개 요인으로 작용할 수 있음을 시사한다. 결과적으로 GDPR은 단순히 법적 규제 장치가 아니라, 플랫폼 구조와 이용자 인식의 경계를 재설정하여 담론 참여의 형태와 수준을 동시적으로 재구성하는 제도적 사건으로 이해될 수 있다.

이와 같은 구조적 긴장 속에서 GDPR은 온라인 참여 구조에 외생적 제도 충격(exogenous institutional shock)으로 작용했다. GDPR은 데이터 수집·처리 기준을 엄격히 강화하면서, 온라인 광고의 효율성을 감소시켰다는 보고(Goldfarb & Tucker, 2011), 웹사이트 트래픽의 재편(Aridor et al., 2020), 그리고 알고리즘 설계의 투명성 제약(Binns, 2018; Johnson et al.,

2020) 등 다양한 결과를 낳았다. 그러나 기존 논의는 거시적 수준의 이용량 변화나 산업적 효과에 집중되어 있으며, 실제 사용자 담론, 특히 댓글과 같은 미시적 상호작용 층위에서 참여의 양적·질적 패턴이 어떻게 달라졌는지에 대한 실증적 연구는 매우 부족하다.

한편, 최근 한국에서도 개인정보보호법(Personal Information Protection Act, PIPA)의 전면 개정, 마이데이터(MyData) 제도의 확산, 그리고 인공지능 윤리 프레임워크 수립 등 데이터 활용과 보호 간 균형을 모색하는 정책 논의가 활발히 전개되고 있다. 이러한 흐름은 데이터 주체의 자기결정권을 강화하면서도 산업적 혁신을 저해하지 않는 제도 설계라는 점에서 GDPR의 철학적 문제의식과 상통한다. 따라서 GDPR은 단순한 유럽형 규제 모델이 아니라, 한국의 정보화정책이 직면한 데이터 거버넌스 전환기를 이해하는 비교적 거울로 기능할 수 있다. 임규건 외(2022)는 기술적 보안체계와 제도적 신뢰가 결합된 한국형 인터넷 거버넌스 모델을 제시하며, 정책·기술·사회가 통합된 접근의 중요성을 강조하였다.

이에 따라 본 연구는 GDPR이 온라인 담론 구조에 미친 변화를 실증적으로 분석함으로써, 한국 정보화정책이 '책임 있는 데이터 활용(responsible data use)'을 구현하기 위해 요구되는 사회기술적 조건을 탐색하고자 한다. 이를 통해 개인정보 규제, 플랫폼 거버넌스, AI 윤리 등 분절되어 온 연구 흐름을 통합적으로 연결하는 분석 틀을 제시한다.

이러한 공백을 메우기에 Reddit은 중요한 분석 장을 제공한다. Reddit은 업보트(upvote) 기반의 집단 평가 메커니즘을 통해 댓글의 가시성과 생존을 결정하며(Gilbert, 2013), 서브레딧 단위의 자율적 중재는 담론의 질적 수준을 유지하는 집단 규범으로 기능한다(Jhaver et al., 2017; Matias, 2019). Chandrasekharan et al.(2018)은 Reddit 제재 데이터를 통해 커뮤니티 규범과 평가 장치가 결합하여 담론의 질을 조형하는 과정을 보여주었다. 또한 최근 연구는 GDPR 시행 이후 온라인 프라이버시 논의의 주제적 구성이 변화했음을 텍스트 마이닝 분석을 통해 보고한 바 있으며(김주영, 2025), 이는 규제가 Reddit 담론의 내용적 차원에도 실질적 영향을 미쳤음을 시사한다.

본 연구는 Downs(1972)의 이슈-관심 주기(issue-attention cycle) 이론을 토대로, GDPR이라는 제도적 충격이 Reddit 프라이버시 담론의 참여 패턴을 어떻게 재편했는지를 실증적으로 탐구한다. 이슈-관심 주기에 따르면, 공공의 관심은 특정 사건을 계기로 급격히 고조되지만, 곧 복잡성과 비용 인식이 심화되면서 실용적이고 압축적인 형태로 전환되고, 궁극적으로는 관심이 쇠퇴하는 순환 과정을 따른다. GDPR은 온라인 커뮤니티에서 이러한 관심 주기를 촉발·재구성하는 외생적 사건으로 기능했을 가능성이 크며, 이는 Reddit과 같은 커뮤니티에서 프라이버시 관련 담론의 참여 규모와 질적 특성을 동시에 변화시켰을 것으로 추론된다.

이를 검증하기 위해 본 연구는 2016년 1월부터 2023년 12월까지 Reddit privacy 서브레딧에서 수집된 게시글(최종 2,673개)과 댓글(최종 38,944개)을 분석 대상으로 삼는다. 연구 설계는 두 단계로 구성된다. 첫째, GDPR 시행 전후의 댓글 수와 질적 지표 평균을 비교하여 규제의 단순 효과를 확인한다. 둘째, 다중회귀분석(OLS)과 로버스트 회귀를 활용하여 댓글 수와 질적 지표 간의 관계가 GDPR 이후 어떻게 변화했는지 검증한다. 이를 통해 본 연구는 참여 규모와 질 간 관계의 제도적 재편을 실증적으로 드러내고자 한다.

학문적 기여는 세 가지로 요약된다. 첫째, Reddit 담론 데이터를 활용하여 참여 연구를 양적 차원에서 질적 차원까지 확장한다. 둘째, GDPR이라는 제도적 사건을 계기로 규제-사용자 담론 간 연결고리를 자연실험적 맥락에서 규명한다. 셋째, 로그 변환과 로버스트 회귀 등 빅데이터 적합적 분석 방법론을 적용하여 극단값과 이분산성을 보정함으로써 참여 구조 재편 메커니즘을 제시한다.

이를 종합해보면 본 연구는 GDPR 시행을 전후로 온라인 프라이버시 담론에서 나타나는 참여 구조의 변화를 분석하고자 한다. 구체적으로 다음과 같은 연구문제

를 제기한다. "GDPR과 같은 개인정보 규제가 온라인 커뮤니티 담론에서 참여의 양적 규모와 질적 차원의 관계를 어떻게 재구성하는가?" 이를 통해 디지털 규제가 단순히 참여의 양적 증감에 영향을 미치는지를 넘어 참여의 형식과 상호작용 구조를 어떻게 재편하는지를 규명하고자 한다.

Ⅱ. 이론적 배경

1. GDPR과 온라인 담론의 이슈-관심 주기

디지털 사회에서 개인정보 보호는 반복적으로 부상하는 핵심 이슈로, 개인의 권리 보호와 기업의 데이터 활용 간의 긴장을 둘러싸고 다양한 사회적·정책적 담론을 형성해왔다. 데이터 경제의 확장 속에서 개인정보는 기업의 경쟁우위를 결정하는 중요한 자산으로 간주되지만(Acquisti et al., 2016), 동시에 감시와 프라이버시 침해에 대한 사회적 우려 역시 증폭되어 왔다(Zuboff, 2019). 이러한 맥락에서 GDPR은 전 세계적으로 가장 포괄적이고 강력한 개인정보 보호 규제로 평가되며(Voigt & Von dem Bussche, 2017), 기업과 개인의 데이터 수집·이용 관행에 중대한 제도적 충격을 가하였다.

이러한 제도적 사건은 단순히 법률적·기술적 변화를 넘어서, 온라인 공론장에서 담론의 촉발과 재편을 야기할 수 있다. 특히 Reddit과 같은 커뮤니티는 규제 변화에 대한 사회적 반응을 자연실험적 조건으로 관찰할 수 있는 장을 제공한다. 온라인 담론에서 특정 사건이나 정책이 관심을 촉발하는 현상은 이미 다양한 영역에서 보고되어 왔다. 예를 들어, 트위터와 같은 소셜미디어는 사회운동이나 정책 이벤트가 발생했을 때 급격한 참여 증가와 담론 확산을 보여주며(Honeycutt & Herring, 2009; Bruns & Burgess, 2015), 댓글과 같은 사용자 참여 행위는 사건의 의미를 구성하고 사회적 해석을 매개하는 역할을 한다(Papacharissi, 2015).

그러나 이러한 관심은 지속적으로 증가하기보다는 특정한 주기적 패턴을 따른다. Downs(1972)가 제시한 이슈-관심 주기는 대중의 관심이 선형적으로 축적되지 않고, 일정한 단계적 변화를 거친다고 설명한다. 먼저 사전 단계(pre-problem stage)에서 소수 전문가 집단이 문제를 인지하고, 문제 인식 단계(alarmed discovery)에서 사건이나 규제가 대중의 급격한 주목을 끌며, 비용 인식 단계(realizing the cost of significant progress)에서 해결의 복잡성과 부담이 부각되고, 관심 쇠퇴 단계(decline of public interest)를 거쳐, 사후 단계(post-problem stage)에서는 이슈가 다른 주제로 이동하거나 잠재화된다. 이후 대중은 다시 새로운 사건을 통해 순환적으로 문제에 관심을 기울인다. 이러한 주기적 패턴은 환경 문제(McComas & Shanahan, 1999), 공중보건(Hilgartner & Bosk, 1988), 디지털 규제(Wiener, 2004) 등 다양한 분야에서 경험적으로 확증되었다.

본 연구는 GDPR을 이러한 이슈-관심 주기의 관점에서 이해하고자 한다. GDPR 시행은 데이터 활용과 프라이버시 보호 간의 긴장을 제도적으로 가시화한 사건으로, 온라인 커뮤니티 참여 양식에 외생적 충격을 제공하였다. Reddit과 같은 프라이버시 커뮤니티에서 GDPR 직후 댓글 수가 증가한 현상은 Downs(1972)가 제시한 "문제 인식과 관심 급등 단계"로 해석될 수 있다. 반면, 이후 댓글 길이가 감소하거나 점수 분포가 재편된 현상은 "관심의 쇠퇴" 혹은 "구조적 전환" 단계에 해당한다. 즉, GDPR은 단순히 관심의 양적 증가만이 아니라, 참여 질적 구조의 재편이라는 변화를 동반했음을 시사한다.

이러한 해석은 GDPR과 같은 규제가 온라인 담론에 미치는 영향이 단순히 참여량의 확대/축소에 국한되지 않고, 참여의 질적 전환 메커니즘을 야기한다는 점에서 중요한 학문적 함의를 지닌다. 즉, 제도적 규제는 담론의 이슈-관심 주기를 재구성하며, 온라인 커뮤니티 참여의 양-질적 균형을 구조적으로 변동시키는 역할을 수행한다. 본 연구는 이러한 관점에서 GDPR 전후 Reddit 프라이버시 담론의 댓글 참여 데이터를 분석함으로써,

규제가 담론 구조에 미친 실질적 영향을 실증적으로 검토한다.

GDPR이 온라인 담론의 양적·질적 구조를 어떻게 변화시킬 수 있는지 이해하기 위해서는, 규제가 개인 인식·플랫폼 구조·사회 규범을 매개로 작동하는 다층적 경로를 고려해야 한다. 본 연구는 세 가지 이론적 관점을 중심으로 이러한 가능성을 논의한다.

첫째, 커뮤니케이션 프라이버시 관리 이론(Communication Privacy Management Theory, Petronio, 2002)은 개인이 온라인 발화를 잠재적 정보 공개 행위로 인식하고, 위험 인식에 따라 공개 범위를 조정한다고 본다. GDPR은 데이터 삭제권과 접근권을 제도화함으로써 이용자의 통제감과 프라이버시 경계를 강화시켰고, 이에 따라 자기검열과 선별적 발화가 증가할 수 있다(Kim & Zo, 2025). 이는 담론의 표현 방식이 장황한 서사에서 간결하고 신중한 형태로 이동하는 토대를 제공한다.

둘째, Evans et al.(2017)에 따르면 기술적 환경은 이용자 행동의 가능성과 제약을 동시에 규정한다(Affordance Theory). GDPR 시행 이후 Reddit은 데이터 저장, 삭제, 보존 기간 등과 관련된 기능을 조정하였고, 이는 이용자의 발화 지속성과 노출 가능성에 대한 인식에 영향을 주었다. 이로 인해 이용자들은 상대적으로 위험이 낮고 효율적인 상호작용 전략을 택하게 될 수 있다.

비슷하게 Scott (2013)은 제도가 규범적·인지적 압력을 통해 행위를 표준화한다고 본다(Institutional Theory). GDPR은 "프라이버시 보호가 사회적으로 옳다"는 규범을 강화하였고, 커뮤니티의 평가 시스템(업보트·점수)은 이러한 규범을 내면화하는 수단으로 기능할 수 있다(Suchman, 1995). 결과적으로 프라이버시 친화적이고 간결한 발화가 커뮤니티 내에서 더 높은 평가를 받을 가능성이 커진다. 예를 들어, GDPR은 (1) 개인의 프라이버시 경계 관리 강화, (2)플랫폼 제도의 재설계, (3)규범적 평가 구조의 내면화를 통해 참여 행위의 양적·질적 패턴을 재조정할 수 있는 제도적 조건을 형성한다. 이러한 이론적 논의는 GDPR 이후 Reddit 담론에서 관찰된 참여 구조의 변화가 단순한 사건 반응이 아니라, 제도–기술–행위가 교차하는 사회기술적 메커니즘의 결과일 가능성을 시사한다.

2. 온라인 플랫폼에서의 참여

온라인 플랫폼에서 사용자 참여(engagement)는 전통적으로 조회수, 좋아요 수, 댓글 수 등과 같은 양적 지표(volume)로 측정되어 왔다. 이러한 지표들은 사용자의 주목도와 활동량을 직관적으로 보여주지만, 담론의 깊이·설득력·정보의 유용성·사회적 가치 등 질적 차원(quality) 을 반드시 내포하지는 않는다. 실제로 단순한 활동량 지표는 참여를 과대평가하거나, 피상적 반응을 심층적 참여로 오인할 위험을 내포한다(Siersdorfer et al., 2010; Napoli, 2011).

비디오 스트리밍 및 추천 플랫폼의 맥락에서는 이러한 양–질 간 긴장이 더욱 두드러진다. YouTube의 대규모 추천 시스템을 분석한 Covington et al.,(2016)은 플랫폼의 핵심 목표가 시청 시간(watch time) 극대화임을 보여주었다. 이러한 알고리즘적 최적화는 체류 시간을 늘리고 참여의 양을 폭발적으로 증가시키지만, 동시에 짧고 자극적인 콘텐츠가 과도하게 부각되는 품질 저하 현상을 야기할 수 있다(Bishop, 2019). 또한 Zhan et al.(2022)은 TikTok과 같은 마이크로 비디오 플랫폼에서 영상 길이(duration bias) 가 추천 모델 학습에 체계적으로 반영되어, 짧은 콘텐츠가 구조적으로 선호되고 이용자의 소비 주기가 단편화된다고 지적하였다.

사용자 행태 연구 또한 이러한 불균형을 뒷받침한다. Kim & Zo(2025)는 비디오 추천 시스템에서 데이터 감시(dataveillance) 인식이 자기검열과 선택적 공개를 유발함을 발견했으며, Kim et al.(2025)은 데이터 활용의 지각된 가치가 낮을수록 사용자가 정보 공개를 중단(discontinuation)하는 경향을 보고하였다. 이러한 결과는 제도적 규제나 알고리즘 설계가 참여 규모와 질적 평가 간의 관계를 재편할 수 있다는 본 연구의 분석 틀

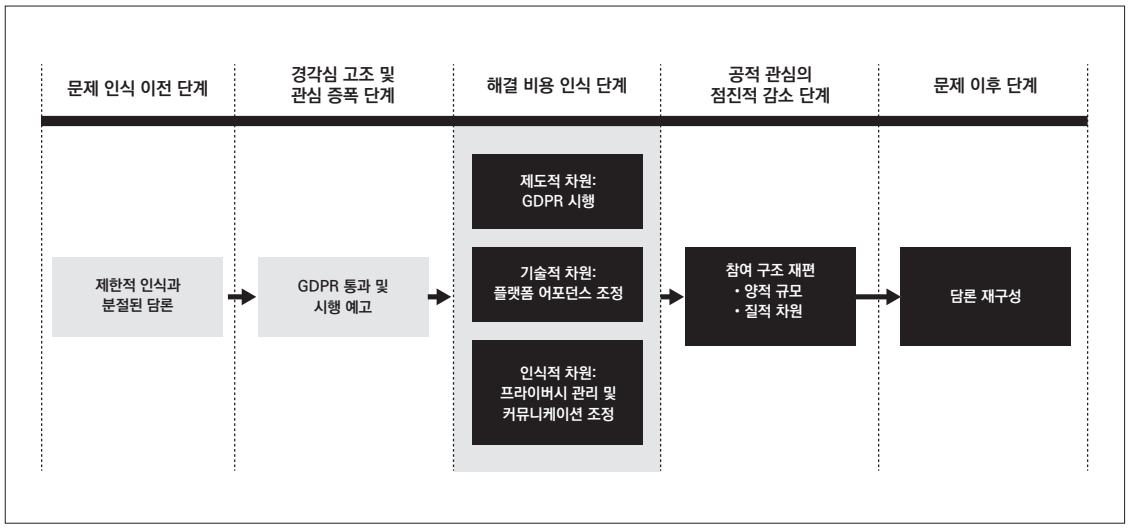
을 뒷받침한다.

댓글과 토론 환경에서는 참여의 질적 차원을 어떻게 정의하고 측정할 것인가가 주요 학문적 과제로 부상하고 있다. 최근 연구들은 댓글 점수(score, 업보트 기반), 댓글 길이(length), 논쟁성(controversy), 정중함(politeness), 근거 제시(argumentation), 감성(emotion) 등 다양한 지표를 활용하여 참여의 질을 다차원적으로 측정하고자 하였다. 예컨대 Berg et al.(2026)는 사용자 생성 콘텐츠의 품질을 정중함 · 논거 · 근거 제시 등으로 평가하며, 단순한 길이나 점수만으로는 담론의 숙의적 가치를 충분히 포착하기 어렵다고 지적했다. Kolhatkar et al.(2020) 또한 뉴스 댓글 연구에서 constructiveness 척도를 개발하여, 댓글이 토론을 심화시키는 정도를 계량화하였다. 이러한 시도들은 온라인 참여를 단순히 수치화된 반응량이 아니라, 사회적으로 유의미한 기여 여부로 평가해야 함을 강조한다.

특히 Reddit은 댓글 참여 연구의 대표적 사례를 제공한다. Reddit의 업보트 기반 평가는 댓글의 가시성과 생존을 결정짓는 핵심 메커니즘이며(Gilbert, 2013), 이는 단순한 개인 의견 반영을 넘어 집단 규범 실행의 장으로 작동한다(Matias, 2019). Jhaver et al.(2017)은 서브레딧(subreddit)의 자율적 중재가 커뮤니티 규범과 결합하여 담론의 질을 유지하는 효과를 보인다고 분석했으며, Chandrasekharan et al.(2018)은 대규모 삭제 · 차단 데이터를 통해 커뮤니티 규범 집행과 평가 구조가 상호작용하여 온라인 토론의 질을 형성하는 과정을 실증적으로 입증했다. 이러한 연구들은 Reddit 댓글 참여가 단순한 양적 반응의 집합이 아니라, 플랫폼 설계와 집단 규범이 교차하는 사회기술적 구조임을 보여준다.

나아가 참여의 양적 · 질적 구분은 오늘날 디지털 정책과 알고리즘 설계 논의에서도 중요한 함의를 지닌다. Napoli(2011)는 "미디어 참여 지표의 양적 편중"이 민주적 숙의와 공적 담론의 질을 약화시킬 수 있음을 지적했으며, Stroud et al.(2017)은 온라인 토론장에서 질적 참여를 강화하기 위한 제도적 · 알고리즘적 개입의 필요성을 제시했다. 이러한 맥락에서 GDPR과 같은 제도적 변화는 단순히 참여량의 변화를 유발하는 수준을 넘어, 양-질 간 전환 메커니즘을 재구성하는 잠재력을 지닌다.

본 연구는 이러한 논의를 Reddit 프라이버시 담론 분석에 적용한다. 특히 Reddit은 댓글이 곧 공적 참여(public engagement)로 기능한다는 점에서 주목할 만하다. Reddit의 업보트 기반 평가 구조는 댓글의 가시성과 생존을 결정하며(Gilbert, 2013), 이는 단순한 개별 의견 표출이 아니라 공적 승인(public endorsement)을 매개하는 집단적 규범의 실행으로 작동한다(Matias, 2019; Papacharissi, 2015). 따라서 댓글 수는 개인의 활동 빈도를 넘어, 커뮤니티 차원의 공적 관심과 숙의적 참여(deliberative engagement)의 강도를 반영하는 지표로 해석될 수 있다(Stroud et al., 2017). 또한 댓글의 점수(score) 및 길이(length)는 발화의 사회적 평가 구조와 표현의 형식적 특성을 함께 포착한다는 점에서, GDPR 이후 나타난 담론의 형식적 재편(formal restructuring)을 측정할 수 있는 유효한 대리지표(proxy)로 기능한다. 이는 단순히 '질적 향상'이 아니라, 규제 환경 변화에 따른 참여 양식의 구조적 전환을 의미한다. GDPR 시행은 이용자에게 새로운 위험 · 비용 요인을 인식시키는 동시에 규범적 기대를 강화하는 제도적 사건으로 작용했을 가능성이 크다. 따라서 본 연구의 참여 질(quality) 지표는 Reddit 플랫폼 내 평가 구조와 발화 형식의 변화(formal restructuring)를 포착하는 데 초점을 두며, 댓글 길이의 단축이나 점수 집중의 강화는 참여의 질적 향상 그 자체보다는 커뮤니티 담론의 형식적 전환(formal shift)을 반영한다고 해석한다. 이는 짧지만 평가 최적화된 상호작용이 증가하거나, 대규모 참여가 더 강력한 승인 신호로 번역되는 구조를 공고화할 수 있음을 의미한다. 즉, 규제 충격 이후 사용자 담론은 단순히 "활동량의 일시적 팽창 후 쇠퇴"로만 설명되지 않고, 참여 구조의 질적 재편이라는 장기적 구조 변화로 나타날 수 있다.



〈그림 1〉 GDPR에 따른 온라인 프라이버시 담론의 이슈-관심 주기 변화 모형

이와 같은 이론적 틀은 기존 플랫폼 책임(platform liability) 혹은 데이터 프라이버시 규제 연구가 주로 법적 의무나 산업 수준의 대응에 초점을 맞춘 것과 달리(Gorwa, 2019; Suzor, 2020), 규제가 담론 수준에서 공적 참여의 형식을 어떻게 변화시키는가에 주목한다. 즉, 본 연구는 GDPR을 정보화정책의 한 유형으로 재개념화하고, 그 미시적 파급효과(플랫폼의 구조 조정, 이용자 인식 변화, 커뮤니티 평가 규범 강화 등)를 실증적으로 규명함으로써, 정보정책 연구의 범위를 행위와 담론의 수준까지 확장하고자 한다. 종합해보면, 본 연구의 맥락에 맞게 온라인 프라이버시 담론의 관심 주기를 구성해보면 다음 〈그림 1〉과 같다.

Ⅲ. 연구방법

1. 데이터 수집 및 전처리

본 연구는 미국의 대표적 온라인 커뮤니티 플랫폼인 Reddit 중 프라이버시 관련 논의가 가장 활발하게 이루어지는 r/privacy 서브레딧을 분석 대상으로 설정하였다. Reddit은 이용자가 관심 주제에 따라 서브레딧(subreddit)을 자유롭게 생성·구독하며 담론을 전개할 수 있는 구조를 갖고 있어, 방대한 텍스트 데이터를 확보할 수 있을 뿐만 아니라 높은 외부 타당성과 다양한 사용자 속성을 반영할 수 있다는 점에서 학문적 분석에 적합하다(Curiskis et al., 2020).

데이터는 Python 기반의 Reddit API(PRAW 라이브러리)를 활용하여 2016년 1월 1일부터 2023년 12월 31일까지 수집하였다. 수집 대상은 'YouTube' 키워드가 포함된 영문 게시글 및 해당 댓글이다. 'YouTube'를 키워드로 선택한 이유는, 이 플랫폼이 전 세계적으로 가장 널리 사용되는 대표적 비디오 추천 서비스로서(2024년 기준 약 2.53억 명의 월간 활성 사용자, data.ai, 2024), 프라이버시 및 보안 관련 논쟁이 빈번하게 제기되는 사례이기 때문이다(e.g., Amaya et al., 2019; Proferes et al., 2021). 특히 사용자들은 시청 기록, 광고 타겟팅, 계정 연동과 같은 기능에서 프라이버시 우려를 반복적으로 제기해 왔으며, 이에 따라 Reddit 내에서 YouTube 관련 담론은 프라이버시 문제의 대표적 사례로서 학문적·실무적 시사점을 제공한다.

GDPR의 공식 시행일은 2018년 5월 25일임에 따라 데이터는 GDPR 시행 전(2016.01-2018.04) 과 시행

후(2018.06-2023.12)로 구분하였다. 분석 시작 시점을 2016년으로 설정한 이유는 이전 논의 내용의 부재성과 더불어 GDPR 법안이 2016년 4월에 공표된 이후 전 세계적으로 관련 논의가 확산되었기 때문이다. 제도 시행 전의 관심 형성기와 시행 후의 안정화 과정을 포괄함으로써, Reddit 담론의 단기·중기 변화를 함께 검토하였다.

초기 수집 결과는 총 2,676건의 게시글과 42,191건의 댓글이었다. 이후 데이터 정제 과정에서, 댓글 본문이 "[deleted]"로 표시된 경우나 작성자가 삭제된 계정의 경우를 제외하였다. 또한 게시글과 댓글을 연결하는 고유 식별자(post_id)를 기준으로 두 데이터셋을 병합한 뒤, 댓글이 없는 게시글(전체의 약 23%)은 제외하였다. 최종적으로 2,062개의 게시글과 38,944개의 댓글이 분석에 활용되었다.

게시글 데이터(posts)의 고유 식별자(id)와 댓글 데이터(comments)의 참조 식별자(post_id)를 기준으로 두 데이터를 병합함으로써, 각 게시글 단위에서 댓글의 수와 성격을 집계할 수 있는 통합 데이터셋을 구축하였다. 이 과정에서 댓글 본문이 삭제된 경우는 빈 문자열로 처리하여 길이 분석에서는 제외하지 않고 반영하였다. 또한 Reddit 데이터 특성상 일부 게시글은 댓글 수가 비정상적으로 폭증하는 극단치를 포함하므로, 왜도(skewness)를 보정하기 위해 댓글 수(n_comments)에 자연로그 변환(log1p)을 적용하였다.

분석 단계에서는 Reddit 댓글 데이터를 기반으로 참여 규모와 참여 질을 측정할 수 있는 지표들을 구축하였다. 우선 참여 규모는 게시글별 실제 관측된 댓글 수(n_comments)를 활용하여 산출하였다. 이는 특정 게시글이 커뮤니티 내에서 어느 정도의 반응과 주목을 이끌어냈는지를 직관적으로 보여주는 핵심 지표로, GDPR 시행 전후의 참여량 변화를 평가하는 기준이 된다. 참여 질은 세 가지 지표를 통해 포착하였다. 첫째, 댓글 점수 평균(cmt_score_mean)을 산출하여 각 논의가 커뮤니티 내에서 얼마나 긍정적으로 수용되었는지를 반영하였다. 둘째, 업보트 비율(cmt_up_ratio_mean)을 활용하여 댓글 평가 구조의 편향성을 검토함으로써, 긍정적 피드백이 집중되는 경향을 확인하였다. 셋째, 댓글 길이(cmt_len_mean)를 측정하여 담론의 서술적 풍부함과 표현의 다양성을 평가하였다. 이는 단순한 활동량을 넘어, 참여의 질적 특성과 담론의 깊이를 보여주는 보완적 지표로 기능한다. 추가적으로, GDPR 시행 전후의 단기·중기적 변화 추세를 파악하기 위해 월별 단위로 게시글당 평균 댓글 수(mean n_comments)와 평균 댓글 점수(mean cmt_score_mean)를 산출하였다.

2. 분석 방법

본 연구는 GDPR 시행 전후 Reddit r/privacy 커뮤니티에서 나타난 참여 양상의 변화를 다각도로 검증하기 위해 단계별 분석을 수행하였다. 우선, 기초 통계 분석을 통해 댓글 수 및 질적 지표(댓글 점수 평균, 댓글 길이 등)의 분포를 확인하고, 두 시기의 평균 차이를 비교하였다. 이러한 차이가 통계적으로 유의한지를 검증하기 위해 독립표본 t-검정(independent t-test)을 적용하였으며, 등분산 가정은 Levene's test를 통해 사전에 점검하였다. 이를 통해 GDPR 시행 전후에 나타난 단순 평균 차이를 확인할 수 있었다. 또한 GDPR 시행 전후의 단기·중기적 변화 추세를 파악하기 위해 보조적 시계열 분석을 수행하였다. 시간에 따른 참여도의 변화를 시각적으로 파악하기 위해 LOESS (Locally Estimated Scatterplot Smoothing)를 적용한 선형 추세 그래프를 작성하였다.

다음으로, GDPR이 댓글 질적 지표에 미친 효과를 통제된 조건에서 분석하기 위해 다중회귀분석(OLS regression)을 수행하였다. 종속변수는 댓글의 질적 특성을 나타내는 지표(cmt_score_mean, cmt_len_mean 등)로 설정하였으며, 독립변수로는 댓글 수(n_comments) 혹은 로그 변환한 값(log_comments)을 사용하였다. 아울러 GDPR 시행 여부를 더미 변수(0 = 시행 이전, 1 = 시행 이후)로 포함하였고, 참여 규모와 제도적 충격 간의 상호작용 효과를 검증하기 위해 n_comments

〈표 1〉 주요 변수의 기술통계

변수	평균	표준편차	최소값	최대값
댓글 수 (n_comments)	12.33	36.13	0	829
댓글 수의 로그값 (log_comments)	1.66	1.25	0	6.72
댓글 점수 평균 (cmt_score_mean)	32.58	180.88	0	2,904

주: GDPR 집단 비율 — 시행 전 11.8%, 시행 후 88.2%.

× GDPR 항을 모형에 추가하였다. 이를 통해 참여 규모가 댓글 평가나 길이에 미치는 영향이 GDPR 이후 어떻게 달라졌는지를 구체적으로 확인할 수 있었다.

그러나 Reddit 데이터는 소수의 게시글이 비정상적으로 많은 댓글을 유발하는 경우가 존재하여 분포의 왜도와 극단치 문제가 발생한다. 이러한 데이터 특성을 보정하기 위해, 본 연구는 Huber-T norm을 활용한 로버스트 회귀(RLM, Robust Linear Model)를 추가적으로 적용하였다. 이는 극단치의 영향을 완화함으로써 추정 결과의 안정성을 확보하기 위한 조치이다. 또한 OLS 추정치에 대해서도 이분산성 가능성을 고려하여 HC3 robust standard error를 병행 보고하였다.

3. 기초통계

회귀분석에 앞서 주요 변수의 분포 특성을 확인하였다. 〈표 1〉은 Reddit r/privacy 커뮤니티에서 수집된 주요 변수의 기술통계량을 제시한다.

분석 대상 게시글은 총 2,062건이며, 평균 댓글 수(n_comments)는 12.33개(SD = 36.13), 로그 변환된 댓글 수(log_comments)의 평균은 1.66(SD = 1.25)으로 확인되었다. 평균 댓글 점수(cmt_score_mean)는 32.58(SD = 180.88)로 분산이 크게 나타나 Reddit 데이터의 과산포(over-dispersion) 특성을 보여준다.

GDPR 시행 전후의 관측 비율은 Pre 11.8%, Post 88.2%로, Reddit 프라이버시 담론의 대부분이 GDPR 시행 이후에 집중되어 있음을 시사한다. 이러한 분포 특성을 반영하여 이후 회귀분석에서는 로그 변환 및 로버스트 회귀 모형을 병행하였다.

Ⅳ. 연구 결과

1. 참여 규모(Volume)의 변화

GDPR 시행 전후의 댓글 수 평균을 비교하기 위해 독립표본 t-검정을 실시하였다(〈표 2〉 참조). 분석 결과, GDPR 이전(Pre-GDPR)의 평균 댓글 수는 17.71(SD = 43.07), 이후(post-GDPR)는 19.07(SD = 50.39)로 나타났으나, 두 집단 간 차이는 통계적으로 유의하지 않았다(t = 0.47, p = .64). 또한 등분산성 검증을 위한 Levene's test 결과 역시 비유의(p = .57)로 나타나, 두 집단 간 분산 차이도 확인되지 않았다.

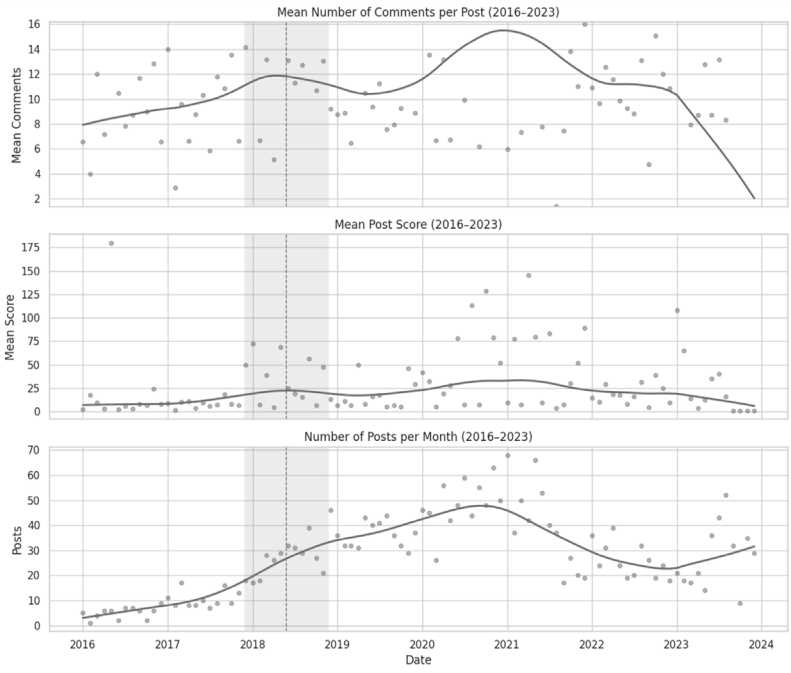
이러한 결과는 GDPR 시행이 Reddit 커뮤니티 내 참여 규모, 즉 댓글 수와 같은 양적 주목도(issue-attention volume)에 직접적인 영향을 미치지 않았음을 시사한다. 다시 말해, 제도적 규제가 이용자들의 '참여량' 자체를 확대하거나 축소하는 효과는 제한적이었다.

그러나 이는 규제가 단순히 참여 규모를 변화시키는 방식이 아니라, 플랫폼의 어포던스 및 행위 전략을 재조정하여 참여의 '형태' 와 '질'을 바꾸는 방향으로 작동했을 가능성을 시사한다(Evans et al., 2017). 즉, GDPR은 발화 빈도 자체보다 이용자가 어떤 방식으로 발화하고 평가에 반응하는지를 구조적으로 재편하는 제도적 맥락으로 작용했을 수 있다.

〈그림 2〉에서 확인할 수 있듯이, 2018년 GDPR 시행 시점을 전후로 댓글 수와 점수가 단기적으로 급등한 뒤

〈표 2〉 GDPR 시행 전후 댓글 수에 대한 기술통계 및 t-검정 결과

집단	평균	표준편차	t	p
시행 전(pre-GDPR)	17.71	43.07	0.47	.64
시행 후(post-GDPR)	19.07	50.39		



〈그림 2〉 이용자 댓글 및 게시글의 월별 추이

점차 감쇠하는 패턴이 나타났다. 게시글 당 평균 댓글 수, 평균 점수, 월별 게시글 수 모두에서 시행 직후 관심 급등이 관찰되었으며, 이후 완만한 하락과 안정화가 이루어졌다. 그래프의 회색 음영은 GDPR 시행일 ±6개월 구간을 나타낸다. 이러한 결과는 Downs (1972)가 제시한 이슈-관심 주기 이론과 일치한다. 즉, 주요 사회적 사건이나 제도 변화 직후 공공의 관심이 단기적으로 폭발적으로 증가하지만, 시간이 지나며 제도화 및 일상화 과정 속에서 점차 관심이 감소하는 '관심 급등–감쇠' 현상이 온라인 프라이버시 담론에서도 동일하게 나타남을 시사한다.

2. 참여 규모(Volume)와 질(Quality), 그리고 평가(Score)의 관계

다음으로, 참여 규모와 댓글 평가(평균 댓글 점수)와의 GDPR 전후 관계를 검증하였다. Reddit 데이터는 일부 게시글에 댓글이 수천 개 이상 몰리는 heavy-tailed distribution을 보이므로, 분석 과정에서 댓글 수는 log(n_comments+1)로 변환하여 사용하였다.

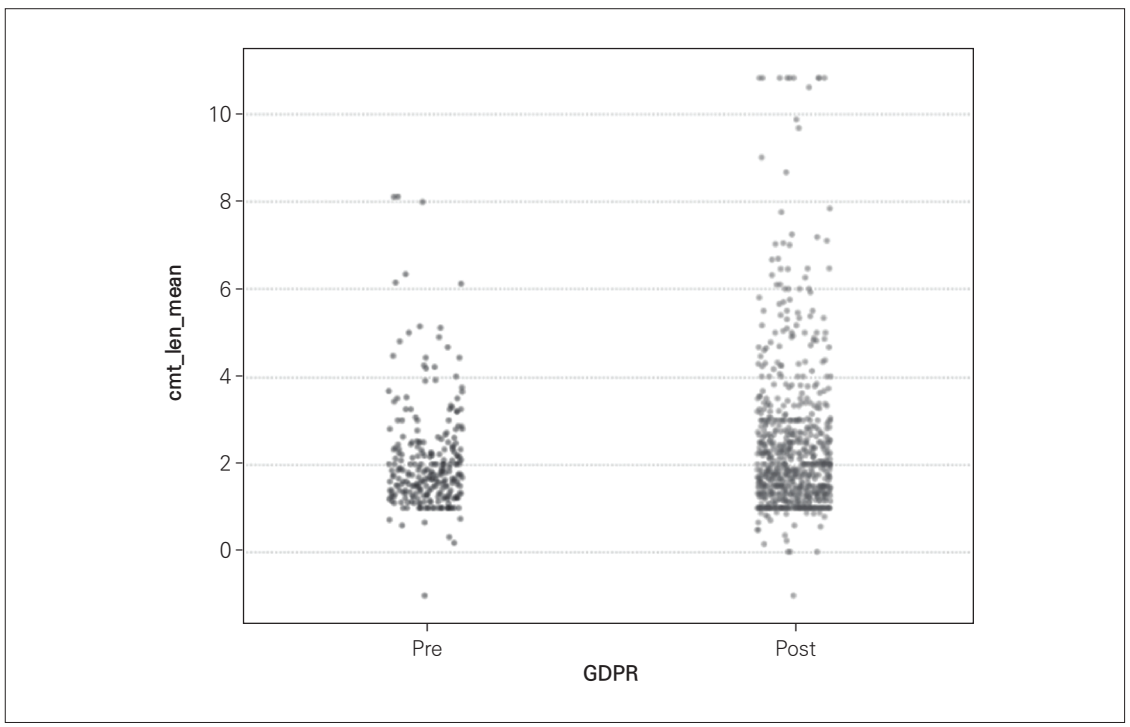
OLS 회귀분석(HC3 robust SE 적용) 결과, log_comments × GDPR 상호작용항은 통계적으로 유의한 양(+)의 효과를 보였다(β = 0.49, p < .01). 즉, GDPR 시행 이후에는 댓글 수가 증가할수록 평균 점수가 더욱 가파르게 상승하는 경향이 나타났다. 동일한 패턴은 로버스트 회귀분석(RLM, Huber 가중치)에서도 확인되었으며, 상호작용항은 역시 유의한 양의 효과를 보였다(β = 0.22, p < .01).

이 두 분석 결과는 〈표 3〉, 〈그림 3〉과 〈그림 4〉에 제시하였다. 먼저 〈그림 1〉에서 확인할 수 있듯이, GDPR

〈표 3〉 댓글 점수에 대한 회귀분석 결과: OLS(HC3)와 로버스트 선형모형(Huber) 비교

독립변수	OLS (HC3) β (SE)	RLM (Huber) β (SE)
절편	0.98*** (0.30)	1.14*** (0.16)
log(댓글 수+1)	0.47** (0.14)	0.33*** (0.07)
GDPR (1=시행 후, 0=시행 전)	−0.71* (0.35)	−0.30† (0.17)
log(댓글 수) × GDPR	0.49** (0.16)	0.22** (0.07)
표본 수 (N)	2062	2062
결정계수 (R^2) / 유사 결정계수	.204	—

주: † p < .10, * p < .05, ** p < .01, *** p < .001.



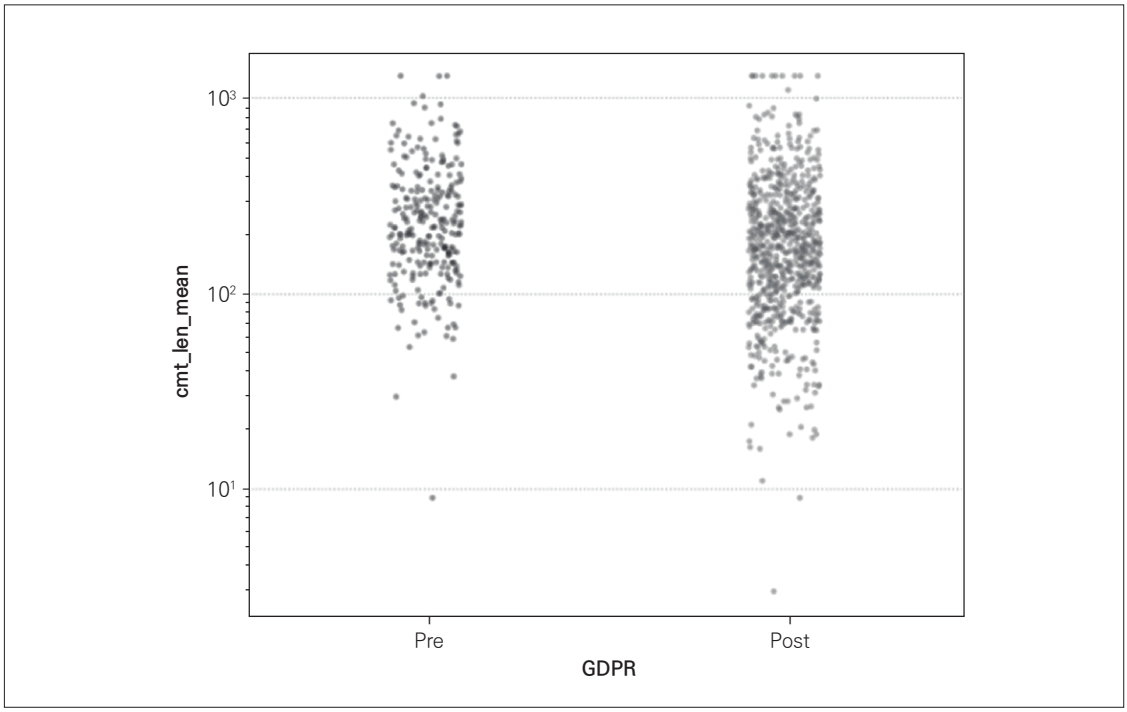
〈그림 3〉 GDPR 시행 전후 댓글 점수의 변화

시행 이후(post-GDPR) 집단은 평균 댓글 점수가 전반적으로 더 높은 방향으로 분포가 이동하였다. 이는 GDPR 이후 커뮤니티가 특정 주제나 게시글에 대해 보다 강한 지지를 표시하는 경향이 강화되었음을 시사한다. 즉, 댓글 수가 늘어나는 상황에서 단순히 양적 증가에 그치지 않고, 긍정적 평가가 집중되는 메커니즘이 제도적 변화 이후 더욱 뚜렷해졌다고 해석할 수 있다. 이러한 양상은 GDPR이 '무엇이 좋은 발화인가'에 대한

〈표 4〉 댓글 수에 대한 음이항 회귀분석 결과

독립변수	β	표준오차	z	p	유의성
절편	1.185	0.107	11.08	< .001	***
log(점수)	0.518	0.049	10.50	< .001	***
GDPR (1=시행 후)	−0.208	0.112	−1.85	.065	†
log(점수) × GDPR	0.068	0.052	1.31	.192	—
로그우도	−8002.1				
표본 수 (N)	2062				
유사 결정계수 (Cox-Snell)	0.669				

주: † p < .10, * p < .05, ** p < .01, *** p < .001.



〈그림 4〉 GDPR 시행 전후 댓글 길이의 변화

집단적 평가 기준을 재구성했을 가능성을 보여준다. 제도이론(Scott, 2013)에 따르면, 규제는 단순한 법적 장치가 아니라 사회적 정당성의 기준을 형성한다. GDPR 시행 이후, 프라이버시 친화적이고 간결한 발화가 더 높은 업보트를 받는 경향은 규범적 승인 구조의 강화(institutionalized reward structure)로 이해할 수 있다. 즉, 제도 변화가 커뮤니티의 평가 규범을 매개하여 '양적 참여 → 질적 집중'으로의 전환을 촉진한 것이다.

Reddit 데이터의 분포가 과산포(over-dispersion) 특성을 보임에 따라, 보조적으로 음이항 회귀모형(Negative Binomial GLM)을 적용하였다. 분석 결과, 댓글 점수(log_score)가 높을수록 댓글 수가 유의하게 증가하였으며(β = 0.52, p < .001), GDPR 시행 변수는 약한 음의 방향성을 보였다(β = −0.21, p = .065). 상호작용항(log_score × GDPR)은 비유의였으나, 계수 방향은 긍정적으로 나타나, GDPR 이후에도 긍정적 평가가 집중될수록 참여가 강화되는 패턴이 유지되었음을 시사한다([표 4]). 이는 Reddit 데이터의 과산포 구조를 보정하더라도, 기존 OLS 및 RLM 결과가 일관된 방향성을 유지함을 보여준다.

반면 〈그림 4〉를 보면 댓글 길이 분포는 로그 스케일 기준으로 GDPR 이후 전반적으로 짧아지는 경향을 보였다. 일부 긴 댓글은 여전히 존재하지만, 평균적으로는 댓글이 간결해지고 분산도 줄어드는 모습이었다. 이는 GDPR 이후 참여 질적 특성이 "심층적 서술"에서 "간결한 반응"으로 이동했음을 보여주며, 폭넓은 담론보다는 짧지만 집약된 상호작용을 통해 주제를 소화하는 쪽으로 구조가 재편되었음을 시사한다.

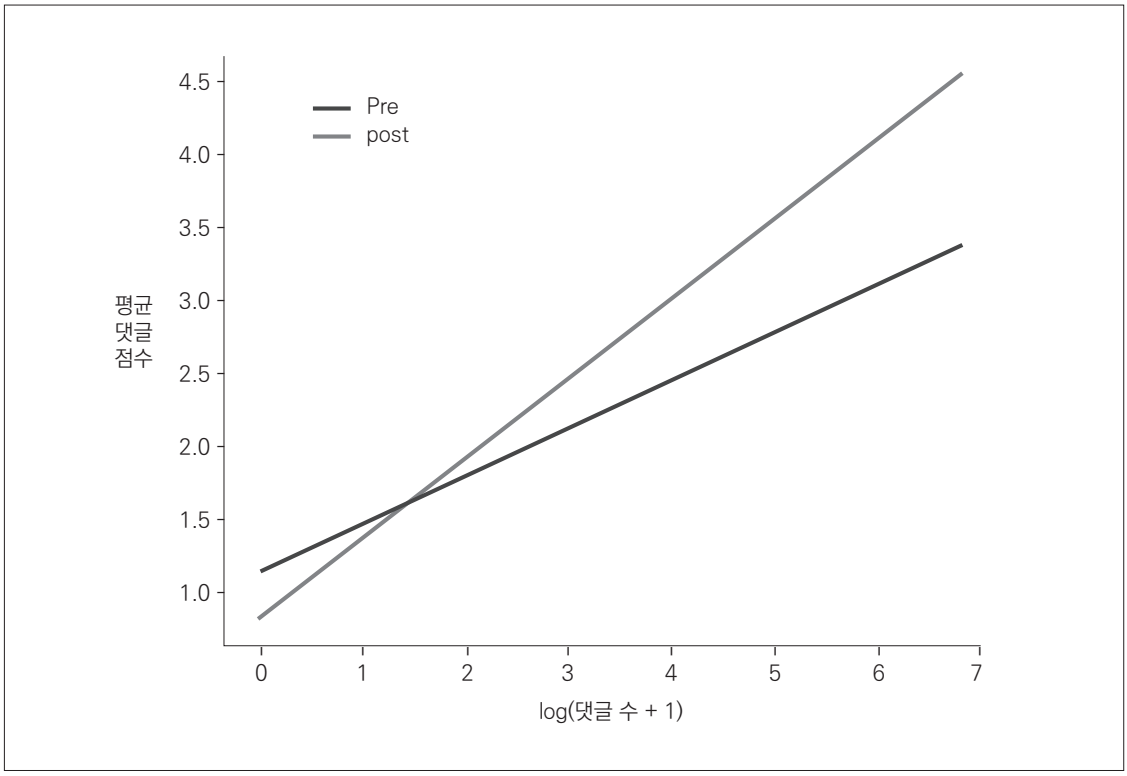
3. 참여 규모(Volume)와 댓글 길이(Length)의 상호작용

댓글 길이를 종속변수로 동일한 모형을 추정한 결과, log(comments) × GDPR 상호작용항은 통계적으로 유의하지 않았다(〈표 5〉). 즉, 댓글 수 증가가 댓글 길이에 미치는 영향은 GDPR 전후에 구조적 차이를 보이지 않았다. 다만, GDPR의 직접 효과는 분석 방법에 따라 상이한 결과가 나타났다. OLS(HC3) 추정에서는 GDPR 이후 댓글 길이가 평균적으로 짧아지는 방향성을 보였으나, 그 효과는 통계적으로 유의하지 않았다(β = −60.85, p = .18). 반면, 로버스트 회귀(RLM, Huber) 분석에서는 GDPR 시행 이후 댓글이 평균 약 60자 감소하는 효과가 유의하게 확인되었다(β = −60.11, p < .01). 이러한 결과는 Reddit 담론이 GDPR 시행 이후 전반적으로 간결화되는 방향으로 재편되었음을 보여준다. 다시 말해, 이전에는 장황하고 설명적인 발화가 상대적으로 많았다면, GDPR 이후에는 실용적이고 핵심적인 정보 전달 중심의 압축적 커뮤니케이션이 강화된 것으로 해석할 수 있다. 이러한 간결화는 커뮤

〈표 5〉 댓글 길이에 대한 회귀분석 결과: OLS(HC3)와 로버스트 선형모형(Huber) 비교

독립변수	OLS (HC3) β (SE)	RLM (Huber) β (SE)
절편	319.75*** (42.10)	230.88*** (21.21)
log(댓글 수+1)	−17.39 (14.50)	3.02 (8.54)
GDPR (1=시행 후, 0=시행 전)	−60.85 (45.42)	−60.11** (22.48)
log(댓글 수) × GDPR	4.94 (15.60)	6.12 (9.04)
표본 수(N)	2062	2062
결정계수 (R^2) / 유사 결정계수	.008	—

주: † p < .10, * p < .05, ** p < .01, *** p < .001.
주: OLS는 HC3 이분산 강건 표준오차를 적용하였으며, 로버스트 선형모형은 Huber 추정량을 사용함.



〈그림 5〉 GDPR 시행 전후 댓글 길이의 분포 비교

니케이션 프라이버시 관리 이론(Petronio, 2002)이 제시하듯, 개인이 정보 공개의 위험을 인식할 때 경계를 강화하고 표현을 최소화하는 '행위적 주의(behavioral vigilance)'의 결과로 볼 수 있다. GDPR은 이용자에게 데이터 공개의 책임과 위험을 재인식시켜, 장황한 설명 대신 요약적 발화와 제한적 정보 공유를 선택하게 한 것이다. 다시 말해, 규제 → 인식 변화 → 발화 전략 변화의 사회심리적 경로가 실증적으로 반영된 결과라 할 수 있다.

〈그림 5〉은 GDPR 시행 전후의 댓글 길이 분포를 비교한 것이다. 전반적으로 post-GDPR 시기에는 댓글이 더 짧은 쪽으로 수렴하는 경향이 뚜렷하며, 긴 댓글의 비중은 줄어드는 양상이 확인된다. 이는 규제 시행 이후 Reddit 커뮤니티에서 담론이 풍부한 서사적 전개보다는 짧고 집약적인 발화를 통해 전개되는 구조로 변화했음을 보여준다.

V. 결 론

본 연구는 GDPR이라는 제도적 충격이 온라인 프라이버시 담론의 참여 양상에 어떠한 변화를 가져왔는지를 Reddit r/privacy 커뮤니티 데이터를 활용하여 실증적으로 규명하였다. 분석 결과, 참여 규모(댓글 수)는 GDPR 전후 통계적으로 유의한 차이를 보이지 않았다. 그러나 참여 규모와 긍정적 평가 간의 관계는 GDPR 이후 유의하게 강화되었으며, 댓글 수가 많은 게시글일수록 평균 댓글 점수가 더 높은 경향이 확인되었다. 반면, 댓글 길이는 전반적으로 짧아지는 추세를 보여, 커뮤니티 담론이 장황한 서사적 발화에서 간결하고 실용적인 정보 중심으로 재편되었음을 시사한다. 보조적으로 수

행한 시계열 분석과 음이항 회귀모형 역시 동일한 방향성을 보여, GDPR 시행 효과가 단순한 평균 차이를 넘어 시간적 리듬과 참여 구조의 질적 전환으로 나타났음을 확인하였다.

이러한 결과는 GDPR이 프라이버시 담론의 양적 확장을 직접적으로 촉진하기보다는, 참여의 질적 패턴을 재편하는 방향으로 작동했음을 보여준다. 즉, GDPR은 규제–플랫폼–행위의 상호작용을 통해 담론의 '형식(form)'을 변화시킨 제도적 매개체로 기능하였다. 구체적으로 첫째, 커뮤니케이션 프라이버시 관리 이론(Petronio, 2002)에 따르면 GDPR은 개인의 정보 공개 경계를 강화하고 자기검열 및 선별적 발화를 촉진함으로써, 담론을 보다 신중하고 핵심 중심적으로 변화시켰다. 둘째, 어포던스 이론(Evans et al., 2017)의 관점에서 Reddit은 GDPR 시행 이후 데이터 삭제 및 보존 기간 명시 등 기술적 설계를 조정하였고, 이용자들은 발화의 지속성과 노출 가능성을 고려한 저위험 · 고효율형 상호작용 전략을 선택하게 되었다. 셋째, 제도이론(Scott, 2013; Suchman, 1995)과 같이 GDPR은 '프라이버시 친화적 발화가 사회적으로 바람직하다'는 규범을 강화하며, 간결하고 책임 있는 발화가 커뮤니티 내에서 더 높은 승인(업보트)을 받는 평가 구조를 정착시켰다. 이 세 가지 경로가 결합되면서 Reddit에서는 참여의 양적 규모보다 질적 집중성이 강화되고, 구성원들은 규제 환경 속에서 간결하지만 사회적으로 승인된 발화를 중심으로 담론을 재조직하게 된 것이다.

이와 같은 결과는 GDPR 규제가 이용자의 인식 · 플랫폼의 기술 구조 · 커뮤니티의 사회적 규범을 매개로 담론 구조를 재편하는 사회기술적 제도로 이해할 수 있음을 보여준다. 다시 말해, GDPR 규제는 데이터 보호의 법적 목적을 넘어 디지털 공론장의 이슈-관심 주기를 변화시키고, 참여의 사회심리적 · 기술적 · 규범적 토대를 재구성하였다. 이는 개인정보보호 연구를 법제 중심에서 행위 · 구조 중심으로 확장하는 새로운 분석 틀을 제시한다.

학문적으로 본 연구는 다음과 같은 독창적 기여를 가진다. 첫째, 기존 개인정보 연구가 주로 법적 · 정책적 효과나 기업 대응에 초점을 맞춘 데 비해, 본 연구는 실제 온라인 참여 데이터를 통해 규제가 개인 행위 수준의 사회심리적 변화를 어떻게 유발하는지를 실증적으로 밝혔다. 이는 Downs(1972)의 이슈-관심 주기 이론을 디지털 맥락에서 확장하며, 규제가 단순한 관심의 급등이나 감쇠를 넘어 참여 규모의 질적 전환 메커니즘을 재편할 수 있음을 보여준다. 둘째, Reddit의 업보트/댓글 구조를 활용하여 참여 질을 다차원적으로 측정하고, GDPR이 이러한 지표에 미친 차별적 효과를 확인하였다. 이를 통해 제도적 변화가 커뮤니티의 소통 방식에 미치는 다면적 영향을 심층적으로 분석한 선행 연구를 보완하였다. 셋째, 제도 변화가 커뮤니티 소통 방식에 미치는 다면적 영향을 분석하는 방법론적 틀을 제공하였다. 이는 플랫폼 거버넌스와 정보화 정책 논의에서, 기술적 설계와 사회적 규범의 교차적 영향력을 함께 강조하고 있다.

실무적으로 본 연구는 다음과 같은 시사점을 제공한다. 첫째, 개인정보 보호 규제가 온라인 담론의 양적 활성화보다는 질적 변화, 특히 응축된 정보 교환과 특정 주제에 대한 집중적 지지 형성으로 이어질 수 있음을 인지하며, 규제 설계 시 이러한 미묘한 커뮤니케이션 재편 효과를 고려해야 한다. 규제는 단순한 법적 의무를 넘어 담론의 양적 · 질적 패턴을 형성하는 사회적 장치임을 보여주며, 향후 개인정보 정책 설계 시 사회적 · 문화적 효과를 고려해야 함을 시사한다. 둘째, 기업들은 GDPR과 같은 규제가 소비자들의 데이터 보호에 대한 인식을 높여, 프라이버시 관련 주제에 대한 심도 깊은 논의 참여를 유도한다는 점을 이해하고, 투명하고 간결한 정보 제공 전략을 통해 신뢰를 구축해야 한다. 이러한 정책의 변화가 소비자 신뢰와 긍정적 참여를 촉진할 수 있는 전략적 기회로 활용될 수 있음을 시사한다. 마지막으로, GDPR 이후 댓글이 간결화된 현상은 커뮤니티 운영자에게 장황한 참여보다 요약 · 핵심 정보 기반의 상호작용을 촉진하는 시스템 설계의 필요성을 제안한다. 요컨대, GDPR은 온라인 프라이버시 담론의 양적 활성화를

직접적으로 촉진하지는 않았으나, 이용자 인식 · 플랫폼 설계 · 집단 규범이 상호작용하는 복합적 경로를 통해 참여의 질적 패턴을 재구성하였다.

나아가 이러한 발견은 유럽 사례를 넘어 한국의 정보화정책에도 함의를 제공한다. 한국의 개인정보보호법, 마이데이터 제도, 인공지능 윤리 프레임워크는 모두 데이터 활용과 보호 간 균형을 모색한다는 점에서 GDPR의 철학적 기반과 맞닿아 있다. 본 연구에서 확인된 '참여 구조의 질적 전환'은, 데이터 주체의 자기결정권 강화와 신뢰 기반 데이터 순환을 중시하는 국내 정책 방향과 일맥상통한다. 특히 규제가 디지털 거버넌스 내에서 참여 양식과 규범을 정교화하는 메커니즘으로 작동할 수 있음을 보여줌으로써, 향후 국내 정보화정책이 설계 · 평가 단계에서 담론 수준의 사회적 효과를 고려해야 함을 시사한다.

또한 플랫폼 책임의 측면에서도, Reddit 사례는 국내 정책 논의와 연결된다. GDPR 시행 이후 Reddit이 데이터 삭제 · 보존 기간 정책을 조정한 것처럼 국내 플랫폼 기업 역시 개인정보 보호위원회의 「온라인 플랫폼 이용자 보호 종합대책(2024)」 등과 연계하여 기술적 설계와 이용자 자율성의 균형을 재구조화할 필요가 있다. 이는 한국의 디지털 신뢰 기반 거버넌스 구축에 있어, 규제-기술-참여 간의 선순환 구조를 모색하는 방향성을 제시한다.

그럼에도 본 연구는 몇 가지 한계를 지닌다. 첫째, Reddit의 단면적 비교에 기반하였기 때문에, 장기적 변화를 완전하게 포착하기는 어렵다. 향후 연구에서는 월별 또는 분기별 시계열 분석이나 segmented regression, difference-in-differences 설계를 통해 제도 충격의 지속성과 재조정 효과를 검증할 필요가 있다. 둘째, 댓글 점수, 업보트 비율, 댓글 길이 등 형식적 지표는 참여의 내용적 질을 직접적으로 측정하지 못한다. 향후 연구에서는 자연어처리 기반의 텍스트 분석을 활용하여 논거성, 정중함, 구성성, 감정 균형 등의 요소를 검증함으로써, 제도 충격이 담론의 의미적 질에 미치는 영향을 분석해야 한다. 셋째, 본 연구는 Reddit r/privacy 커뮤니티를 단일 사례로 다루었기 때문에, 동기간 플랫폼 내부 정책 조정이나 외부 사건 등의 영향을 완전히 배제하기 어렵다. 향후 연구에서는 비규제 커뮤니티를 대조군으로 설정한 접근을 통해 인과적 타당성을 강화해야 할 것이다.

■ 참고문헌

개인정보보호위원회 (2024). 〈온라인 플랫폼 이용자 보호 종합대책〉. 서울: 개인정보 보호위원회.

김주영 (2025). 온라인 비디오 추천 시스템의 프라이버시 논의 동향: GDPR 도입 전후 텍스트 마이닝을 중심으로. 〈윤리경영연구〉, 25권 1호, 91-110. https://doi.org/10.34273/kjbe.2025.25.1.004

다니엘 · 페레즈 (2020). The effects of GDPR on the digital economy: Evidence from the literature. 〈정보화정책〉, 27권 3호, 3-18.

이창범 (2021). 개인정보 이동권의 법적 이슈와 입법 정책 방향. 〈정보화정책〉, 28권 4호, 54-75.

임규건 · 김해연 · 안재익 (2022). 사이버 보안 강화를 위한 한국형 미래 인터넷 추진 방안에 관한 연구. 〈정보화정책〉, 29권 1호, 24-37.

Acquisti, A., Taylor, C. & Wagman, L. (2016). The economics of privacy. *Journal of Economic Literature, 54*(2), 442-492. https://doi.org/10.1257/jel.54.2.442

Berg, J., Fagerholm, A. & Strandberg, K. (2026). Quality user-generated content? A case study of the quality of online news comments on the site of Finnish public service broadcaster Yle. *Journalism Practice, 20*(1), 338-359. https://doi.org/10.1080/17512786.2024.2329759

Bishop, S. (2019). Managing visibility on YouTube through algorithmic gossip. *New Media & Society, 21*(11-12), 2589-2606. https://doi.org/10.1177/1461444819854731

Bruns, A. & Burgess, J. (2015). *Twitter hashtags from ad hoc to calculated publics. In N. Rambukkana (Ed.)*, Hashtag publics (pp. 13-28). Peter Lang.

Chandrasekharan, E., Pavalanathan, U., Srinivasan, A., Glynn, A., Eisenstein, J. & Gilbert, E. (2018). *You*

can't stay here: The efficacy of Reddit's 2015 ban examined through hate speech. Proceedings of the ACM on Human-Computer Interaction, 2(CSCW), 1–22. https://doi.org/10.1145/3134666

Covington, P., Adams, J. & Sargin, E. (2016). *Deep neural networks for YouTube recommendations.* Proceedings of the 10th ACM Conference on Recommender Systems, 191–198. https://doi.org/10.1145/2959100.2959190

Curiskis, S. A., Drake, B., Osborn, T. R. & Kennedy, P. J. (2020). An evaluation of document clustering and topic modelling in two online social networks: Twitter and Reddit. *Information Processing & Management, 57*(2), 102034. https://doi.org/10.1016/j.ipm.2019.04.002

Downs, A. (1972). Up and down with ecology: The "issue-attention cycle." The *Public Interest, 28*, 38–50.

Evans, S. K., Pearce, K. E., Vitak, J. & Treem, J. W. (2017). Explicating affordances: A conceptual framework for understanding affordances in communication research. *Journal of Computer-Mediated Communication, 22*(1), 35–52.

Ganglmair, B., Krämer, J. & Gambato, J. (2024). Regulatory compliance with limited enforceability: Evidence from privacy policies. *ZEW Discussion Paper*, 24-012.

Gilbert, E. (2013). *Widespread underprovision on Reddit*. Proceedings of the 2013 Conference on Computer Supported Cooperative Work, 803–808. https://doi.org/10.1145/2441776.2441866

Goldfarb, A. & Tucker, C. (2011). Advertising bans and the substitutability of online and offline advertising. *Journal of Marketing Research, 48*(2), 207-227. https://doi.org/10.1509/jmkr.48.2.207

Goldberg, S., Johnson, G. & Shriver, S. (2019). Regulating privacy online: The early impact of the GDPR on European web traffic & e-commerce outcomes. SSRN. https://doi.org/10.2139/ssrn.3421731

Gorwa, R. (2019). What is platform governance?. *Information, Communication & Society, 22*(6), 854–871. https://doi.org/10.1080/1369118X.2019.1573914

Hilgartner, S. & Bosk, C. L. (1988). The rise and fall of social problems: A public arenas model. *American Journal of Sociology, 94*(1), 53–78. https://doi.org/10.1086/228951

Honeycutt, C. & Herring, S. C. (2009). *Beyond microblogging: Conversation and collaboration via Twitter*. Proceedings of the 42nd Hawaii International Conference on System Sciences, 1–10. https://doi.org/10.1109/HICSS.2009.89

Jhaver, S., Vora, P. & Bruckman, A. (2017). *Designing for civil conversations: Lessons learned from Reddit*. Proceedings of the ACM on Human-Computer Interaction, 1(CSCW), 1–25. https://doi.org/10.1145/3134697

Kim, J. & Zo, H. (2025). Am I watching or being watched? Exploring the selective disclosure paradox in users' self-censorship to dataveillance awareness in video recommender systems. *Telematics and Informatics*, 102253. https://doi.org/10.1016/j.tele.2025.102253

Kim, J., Zo, H. & Jun, J. (2025). How dataveillance shapes user behavior: The role of perceived value in disclosure and discontinuation. *Computers in Human Behavior*, 168, 108655. https://doi.org/10.1016/j.chb.2025.108655

Kolhatkar, V., Thain, N., Sorensen, J. & Taboada, M. (2020). *Classifying constructive comments*. Proceedings of the AAAI Conference on Artificial Intelligence, 34(5), 8015–8022. https://doi.org/10.48550/arXiv.2004.05476

Matias, J. N. (2019). *Preventing harassment and increasing group participation through social norms in 2,190 online science discussions*. Proceedings of the National Academy of Sciences, 116(20), 9785–9789. https://doi.org/10.1073/pnas.1813486116

McComas, K. & Shanahan, J. (1999). Telling stories about global climate change. *Communication Research, 26*(1), 30–57. https://doi.org/10.1177/009365099026001003

Napoli, P. M. (2011). *Audience evolution: New technologies and the transformation of media audiences*. Columbia University Press.

Papacharissi, Z. (2015). *Affective publics: Sentiment, technology, and politics*. Oxford University Press.

Parsons, J., Schrider, M., Ogunlela, O. & Ghanavati, S. (2023). *Understanding developers' privacy concerns through Reddit thread analysis.* arXiv preprint, arXiv:2304.07650. https://arxiv.org/abs/2304.07650

Petronio, S. (2002). *Boundaries of privacy: Dialectics of disclosure*. SUNY Press.

Rughiniș, R., Rughiniș, C., Vulpe, S. N. & Rosner, D. (2021). From social netizens to data citizens: Variations of GDPR awareness in 28 European countries. *Computer Law & Security Review, 42*, 105585. https://doi.org/10.1016/j.clsr.2021.105585

Scott, W. R. (2013). Institutions and organizations: Ideas, interests, and identities. Sage.

Siersdorfer, S., Chelaru, S., Nejdl, W. & San Pedro, J. (2010). *How useful are your comments? Analyzing and predicting YouTube comments and comment ratings.* Proceedings of the 19th International Conference on World Wide Web, 891–900. https://doi.org/10.1145/1772690.1772781

Stroud, N. J., Scacco, J. M., Muddiman, A. & Curry, A. L. (2015). Changing deliberative norms on news organizations' Facebook sites. *Journal of Computer-Mediated Communication, 20*(2), 188–203. https://doi.org/10.1111/jcc4.12104

Suchman, M. C. (1995). Managing legitimacy: Strategic and institutional approaches. *Academy of Management Review, 20*(3), 571–610. https://doi.org/10.5465/amr.1995.9508080331

Suzor, N. P. (2019). *Lawless: The secret rules that govern our digital lives.* Cambridge University Press.

Voigt, P. & Von dem Bussche, A. (2017). The EU General Data Protection Regulation (GDPR): A practical guide. Springer. https://doi.org/10.1007/978-3-319-57959-7

Wiener, J. B. (2004). The regulation of technology, and the technology of regulation. *Technology in Society, 26*(2–3), 483–500. https://doi.org/10.1016/j.techsoc.2004.01.033

Zhan, R., Pei, C., Su, Q., Wen, J., Wang, X., Mu, G., ... & Gai, K. (2022). *Deconfounding duration bias in watch-time prediction for video recommendation*. In Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (pp. 4472–4481). ACM. https://doi.org/10.1145/3534678.3539321

Zhou, R., Khemmarat, S. & Gao, L. (2010). *The impact of YouTube recommendation system on video views.* Proceedings of the 10th ACM SIGCOMM Conference on Internet Measurement, 404–410. https://doi.org/10.1145/1879141.1879190

Zuboff, S. (2019). *The age of surveillance capitalism.* PublicAffairs. https://doi.org/10.1145/1879141.1879193